# Reading a scientific paper for fun and profit

Jesse Brunner

2022-08-27

## Step 1: Why are you reading this paper?

What do you want to learn from it? Answers include: * I just want to get an overview of a subject, so the intro might be the most important part. * I'm looking for references to other papers that might be more helpful. * I just need to see the results for this on particular comparison; the rest are not important to me. * I'm trying to see how an introduction or discussion can organized as an example. You may well learn things you were not expecting to learn, which is great! But you need to know where you starting from, otherwise you're reading passively, and that is a recipe for learning very little.

It is not a bad idea to write down your purpose or question(s) before you get bogged down in the details! Indeed, it can keep you from wasting a lot of time or losing focus. Moreover, having a sheet of paper[1] nearby is very handy for writing down notes or questions about words, methods, results, or whatever that you need to come back to later. Trust me, your notes will be a lot more helpful than your recollection!

[1] Paper is key. Writing and drawing by hand eats electronic notes for breakfast! And don't get me started on highlighting…

## Step 2: Skim

Give the paper a quick overview. How is it organized? Are there boxes or figures or conceptual diagrams that might help? How are the results organized? Are there one or many questions, one or many studies? Get a lay of the land before you delve in.

## Step 3: Read the sections slowly

Read each (relevant) section slowly, looking up each paragraph to think about what was just said. If you do not understand, was it because you don't understand a term or phrase? If so, try to look it up or otherwise sort it out (e.g., from context, asking your TA). If it was because you maybe lost attention part way through, try reading it aloud, slowly, with *feeling*. You'd be amazed at how that can clarify things! If it is a really important paragraph, you might need to re-read it several times to understand, but sometimes you just need to get the gist to get to the important stuff.

Within each section you probably want to look for different sorts of information.

## Introduction

- *Why* was this study done? What was the context for it? Is there a controversy they were trying to settle, an idea they thought was right or wrong that they wanted to test, or was there some rate or value they were trying to quantify? Understanding this rationale will be very helpful as you try to piece together the larger context of the field; in a way, they are doing a lot of the hard work for you, so pay attention!
- What question were they trying to ask? What was their hypothesis[2]?
  - One important note is that sometimes a study is less interested in testing an idea, but in seeing "what if." Mathematical models, for instance, are often used to see the consequences of certain assumptions (e.g., what would happen if we introduced a vaccine that worked like this? Or can predators control their prey populations? If so, under what conditions?). These are really useful, but they are not really evidence, in and of themselves. They *can* be seen as evidence when their results are compared with the real world (e.g., "Our results suggest we would need ten times as many predators as prey, which seems exceedingly unlikely under natural conditions…") or intuition (e.g., "Our results illustrate that stress-induced immunosuppression is capable of producing the end-of-semester peak in illness.").
- What kind of study is this? Most of the time you can discern whether they are presenting results from an experiment, a correlational or observational study, a survey, a mathematical model, or

[2] Note sometimes you will see *predictions* stated clearly, but the hypothesis is only really implied. What is the difference? Think of a sentence like, "If such and such is true, I would expect this thing to happen." The first part, where we say how we think the world works, is the hypothesis. It is independent of the particular study or experiment. There could be lots of ways of testing this hypothesis. The second part is a prediction or expectation based on the hypothesis, in the context of this study. They are related, but not the same.

## Methods

Unless the details of how a study were done are really important to your purpose, most of the time you can just get a high-level view of what was done. * Start by trying to sort out what kind of study it was, and then you can see if the basic aspects of it seem reasonable. Was it a survey? Who was being surveyed? Was it a representative or appropriate group? Was it done in a way that seems prone to bias? Was it an experiment? What were the treatments? Was there an appropriate control? * Whatever the study, focus on what comparisons they are making (or hoping to make). This is they key part you need to take from the methods section. * Also pay attention to which group(s) they are studying… While it can be a more nuanced question, you can sometimes spot shortcomings that will be important when it comes to interpreting their data. For instance, if a study involves only college freshman, perhaps we should be skeptical as they try to stretch their results to say something about human behavior in general. Similarly, a lab study on tick-host interactions might be really great, but we should pause for a beat before saying what happens in lab mice–the product of generations of inbreeding and selection for docility, white coats, and more—applies directly to small mammals in the wild.

## Results

Let me start by saying, **do not get hung up on the statistical tests or *P*-values**. Let us just 1) assume they got their stats right and 2) recall that a *P*-value is mot a measure of importance, real-ness, or anything similar. Instead, let us focus on the *effect size* and *precision* of the estimate of effect size. * Effect size is a measure of how much of an effect was observed in response to, say, an experimental treatment or, similarly, a difference in groups in a survey or observational study. Examples would include, "Blood pressure was reduced by 10%[3] in the treatment group relative to the control group in our randomized control-treatment experiment," or "Outbred populations had, on average, 1.2[4] fewer parasites species in them than matched inbred populations." Effect sizes can also be expressed along the lines of a slope, as in "the odds of heart disease increased 1.2-fold with every additional encounter with police." * The precision of an estimate is the inverse of the uncertainty in the estimate, often expressed as standard errors (SE or Ste[5].) or confidence intervals (CI, almost always a 95% CI). As a rule of thumb, a confidence interval is usually a bit less than 2 standard errors. The point is not to get back into the statistics[6], but rather to get a sense of the range of estimates that are consistent with the data. For instance, if a study had an effect size of 10% reduction in blood pressure, but the 95% CI was -2% to 22%, their data are as consistent with the drug having *no* effect as it is with causing a ~20% reduction in blood pressure; there is a good bit of uncertainty in this estimate. Contrast that with a similar study with a 95% CI of 8–12%. * Look for graphs. A good graph, when understood, will tell you the story of their data. It will show you the comparison you want to see. It will give a bit of context. So spend the time making sure you understand what it is telling you. Be sure you can explain what it is saying in plain English. * Keep your eye on the prize: Remember what comparison(s) were most relevant to you and look for those. So even if there is a cool graph that you find really compelling, it is not useful if it does not tell you about the part of the study that important to you.

[3] A relative expression of the difference or effect size can be a proportion, percentage, or fold-change. Be careful not to get confused.

[4] Absolute differences will usually have units—here, parasite species—but also include percentage points. Again, be careful you understand what is being expressed.

[5] You may sometimes see a standard deviation (sd). The standard deviation describes the spread in the data. The standard error describes the spread in estimates (as in, if you could repeat the study lots of times and estimate the value of interest, what would the spread in those estimates look like?

[6] Even practicing scientists get the precise meaning of these quantities all sorts of wrong.

## Discussion

Here is where the results are drawn out and put into context. This section can help you understand what you were just looking at in terms of figures or numbers, so contrast you understanding of the results with what they say about them as a reality check.[7]

Note that discussions also take the study's finding and often connect them to a *lot* of different prior studies and ideas that have been lying around. This is, after all, the chance for the authors to explain where their work fits and, especially, why it is important. Don't get lost in this blizzard of tangents and comparisons with existing work; focus on what you need from the paper.

[7] Do not worry too much if you came to a different conclusion than they did or you think they showed a minor effect while they are talking about it as if it were hugely important. Reasonable people can look at the same data and come to different conclusions. You may well be more right than they were. But still, check back with the results to see if you missed something!

## Step 4: Reflect and write down your thoughts

Now is the time to look up and think, what did I learn from reading this paper? Write down those lessons or your impressions while they are fresh. **Write them in your own words**[8]. You had a purpose when you started reading this paper, yes? Look back and see if you answered your question(s) or sorted out whatever needed sorting. See if you think they brought strong or weak evidence to the party. Consider if you agree with their conclusions or remain skeptical; in either case, why?

[8] You do not need to recap the results; you have the paper for that if you need those details. Instead, what do *you* now know?

I like to write a few sentences on the first page of the printed paper that describe the take-home messages. These notes end up being incredibly useful when I go to actually write up my own work. They convey not just what was found, but how much confidence I have in it, any caveats I thought were important, and often how it relates to other work. For instance, "Set out to test the relationship between virulence and transmissibility in *Daphnia*-metch system, but found they are largely unrelated. Enough replication it seems like a strong conclusion, but conducted under really weird conditions, so I am less certain." See how that gives me the context to remember and understand the paper? It puts me back into the mindset I was in when I read it the first time, which makes it much easier to work it into my current paper I'm writing now[9].

[9] Or would be, if this were not hypothetical.